Journal of Energy Technologies and Policy
ISSN 2224-3232 (Paper) ISSN 2225-0573 (Online)
Vol.4, No.4, 2014
www.iiste.org

Medium-Term Load Forecasting Of Covenant University Using The Regression Analysis Methods

Isaac A. Samuel[1]*, Felly-Njoku Chihurumanya F.*, Adewale Adeyinka A.* and Ayokunle A. Awelewa*
*Department of Electrical and Information Engineering, School of Engineering, Covenant University, Canaan land, P.M.B 1023, Ota, Ogun State, Nigeria
* *E-mail of the corresponding author: isaac.samuel@covenantuniversity.edu.ng*

Abstract

Electric load forecasting is principal to the economic and efficient provision of electric power to meet various load demands for a specified period of time. The focus of this paper is on the medium-term load forecast of Covenant University, a Nigerian tertiary educational institution, using the methods of regression analysis. The hourly load information collected from the University's electric power substation is employed as the sample data for the forecast. Three models based on the linear, compound-growth, and cubic methods of regression analysis are developed, and the load forecast results obtained from these models are compared using the mean absolute percentage error (MAPE) and root mean square error (RMSE) performance metrics. It is observed that, among these three models, the linear model has the best error margin: 0.5792 and 41.34 for MAPE and RMSE, respectively.

Keywords: *load forecasting, medium-term, Covenant University, regression-based methods, MAPE, RMSE.*

1. Introduction

Forecasting is the process of making statements about events whose actual outcomes have not yet been observed [1]. It is the basic facet of decision making [2]. Load forecasting is the projection of electrical load that will be required by a certain geographical area with the use of previous electrical load usage in the said geographical area. Electrical load forecasting has a lot of applications. They include: Energy purchasing and generation, load switching, contract evaluation and infrastructure development. It is a very essential part of an efficient power system planning and operation. This is why it has become a major research area in the field of electrical engineering. The time period in which the forecast is carried out is fundamental to the results and use of the forecast. Short-term forecast, which is from one hour to one week, helps to provide a great saving potential for economic and secured operation of power system; medium-term forecast, which is from a week to a year, concerns with scheduling of fuel supply and maintenance operation; and long-term forecast, which is from a year upwards, is useful for planning operations.

All the classes of load forecasting make use of different models to meet the entire specific objectives of the application. These methods have been further defined according to their different uses. Short-term forecasts are used to ensure the stability of a system. Medium (mid-term) forecasts for generation optimization and long term forecasts for investment planning [3]. All the classes of load forecasting make use of different models to meet all the specific objectives of the application [4]. It is important to note that when it comes to forecasting load, the different kinds of characteristics related to electrical load have to be considered and this makes load forecasting a rather difficult task.

In [5] the three different models used to perform load forecasting based on multi-variable regression and the models were applied to the hourly loads of the Jordanian power system for different years and then the results were compared. And in [6] three models used were based on regression analysis to analyze the electric load capacity of a fast growing urban city (Ikorodu, located in Nigeria) and to estimate its future consumption. The models also were compared It was also one of the approaches used by Aslan *et al.* to predict the long term Electric Peak load forecast of Kuthaya, a town in Turkey [7]. A Non-Linear Regression model was created for mid-term load forecasting to improve the modeling of seasonality [8]. Regression analysis can also be combined with Artificial Neural Networks to create a hybrid model [9] [10] [11].

The case study is Covenant University, one of the first privately-owned universities in Nigeria.. It is located in Canaan land, km 10 Idiroko expressway, Ota, Ogun state, Nigeria. Covenant University is a fast developing community known for its in-depth research work.

This study will show the results carried out from a medium-term load forecast that was carried out using values from August to October 2012 using the regression analysis method. The results derived from this forecast were also compared with the actual values recorded within the period that was forecasted (November to January 2013) and the accuracy of the forecasting results tested using the MAPE and RMSE tools.

2. Load forecasting

Electrical load forecasting is usually made by constructing models. These constructed models are based on relative information and previous load demand data. All forecasts should be based on previous data. Therefore, an electrical load forecast should be based on the previous load data of the region to be forecasted.

From 1973 till 1979, a research team was put together to carry out an independent review of past and present load forecasting methodologies based on the state of Iowa in the United States of America. The team was able divide electricity demand into two forms and they are:

1) Energy demand which is measured in kilowatt-hours (or megawatt-hours).
2) Peak demand which is measured kilowatts (or megawatts).

According to the team, Electric energy demand is very important in the corporate and financial planning divisions of utilities. Peak demand forecasting is usually carried out by engineers or planners in the system planning divisions [12].

There are different methods and models used in load forecasting. Among them are: regression analysis methods, Gaussian Process models, time series, static state estimation method, fuzzy logic and the use of Artificial Intelligence (AI) [13].

A group of researchers were able to divide load forecasting methods into two categories namely:

a) Classical approaches: Approaches that are based on statistical methods and forecast future value of a variable by using a mathematical combination of the historic information [14].
b) ANN based techniques: Techniques that make use of the Artificial Neural Networks [7].

From [15], forecasting methods were divided into two. They are: *qualitative* and *quantitative* methods. Qualitative forecasting methods generally use the opinions of experts to predict future load subjectively. These methods are useful when there is little or no historical data available. These methods include: subjective curve fitting, Delphi method and technological comparisons. Quantitative methods include: regression analysis, decomposition methods, exponential smoothing, and the Box- Jenkins methodology. Forecasts helps in the making of financial decisions, the creation of a system design, the development of fuels and energy policies, for environmental impact studies, rate setting and capacity plans decisions. It aids in planning by improving the decisions made by the planners. The basic boundaries of a forecast model are based on the basic boundaries of the modeler's understanding. Modeling has the potential to help formalize the decision problems, extend the scope of the analysis, and guide the decision process [16].Load forecasting helps an electric utility to make important decisions including decisions on purchasing and generating electric power, load switching, and infrastructure development [15].

3. Methods and implementation of the forecast

Regression Analysis is a very important statistical tool which is very useful in determining the statistical relationship or dependence between a change in one variable and the change in another when compared to it. To carry out a medium term forecast of the electrical load consumed by Covenant University, the following steps have to be taken using the collated load data:

3.1 Collection of data:

The data was collected on an hourly basis from the substation, the duration of one year January 2012- January 2013. The load data from August to October 2012 is will be used for the load forecast. It was then transferred and saved as a Microsoft Excel worksheet.

3.2 Pre-processing of the Data Collected

The data collected is then analyzed and arranged into days. The average load consumption per day is calculated and then the total average monthly load consumption from August to October 2012 is also derived. All the calculations were carried out using Excel functions necessary for the data collation. The data was also placed in tables for easier display and comparison.

3.3 Building of the Forecast Models

To carry out medium term load forecasting, there are a lot of models available for use. These models incorporate various data like time, temperature, load and the effect of seasonality. Regardless of the complexity or simplicity of any model, it is important to choose the model that best suits the data available for forecasting.

The data gathered is peculiar to Covenant University, Nigeria. After representing the data in form of graphs, trends were discovered and these trends were used for forecasting. The trends were as a result of observations made from the gathered data. It was noted that:

1) During the hot season, there was a high increase in the load consumed by both the mission and

Covenant University. Further research led to the discovery that a lot of Air conditioners and fans were on to reduce the effect of the heat on homes, classes and shops. The reverse was the case for the cold season as there are few homes in Nigeria with in-built heaters in them, a lot of people will resort to switching off air conditioners and fans, thereby leading to a reduction of load consumption.

2) On weekdays, the total load consumed by Canaan land as a whole was always higher than the total load consumed on weekends.
3) A higher proportion of load was consumed by the Mission when the university was on holiday. This means that there was a reduction of human and industrial activity in the university especially around the month of December and the early days in January.
4) A summary of the patterns above led to a hypothetical grouping of daily time periods into two: Peak period and Non-peak period. Peak period was discovered to be the period between 6pm to 9pm in the evenings when there was a lot of activity in the residential areas e.g. lecturers' quarters and hall of residence and Non-peak period was the period when people were at work. Most of the load consumed was centered on the educational areas of the university e.g. the departments, lecture theatres and administrative areas.
5) When studying the values of load consumption with respect to the peak period, non-peak periods and the different seasons, it was observed that the hot season extends the peak period from 9pm to around 12am. This was because of the extended usage of fans and air conditioners in the residential areas.

The constant variations in the load values caused by the unavailability of power at different points in time due to the erratic power supply by PHCN made the normalization of the data very necessary.

3.4 Carrying out the forecast

This is where the load values that have been accumulated and processed, then substituted into the different regression models that have been built and then solved mathematically. In this process, the accuracy of the forecast values are also tested using the MAPE and RMSE methods.

3.5 Implementation of the Models

It is important to note that the electric load is the most important input needed when carrying out an electrical load forecast. Due to wrong readings of the values and general human error, some load values were either omitted or out of range. These values were corrected using interpolation or by finding the average of the preceding and succeeding values in the daily load log.

The average load consumption for each day is then calculated and added to the other values of average load consumption for the other days in a month to give the average load consumption for the particular month. The values for all the six months were collated and their respective average load curves plotted but to prepare the forecast models, the months of August, September and October 2012 are used. The values for these months are arranged and placed in Table 1 in the column labelled 'Y (load) [in MW]'. The months are in the column 'X (month)'. In this table, August 2012 =1, September 2012 = 2, October 2012 =3.

The Three models are used to calculate the medium term load forecast. They are:

1) The Linear Regression Model: This is the simplest model of regression analysis. It comprises of an independent and dependent variable. It is denoted by the formula:

$$Y = a + bX \tag{3a}$$

Where

Y is the value of the load

X is the number of months (i.e.) August = 1, September= 2 etc.

The values a and b are derived by solving (3b) and (3c) simultaneously.

$$\sum Y = n\,a + b\sum X \tag{3b}$$

$$\sum XY = a\sum X + b\sum X^2 \tag{3c}$$

To carry out the forecasts, the months to be forecasted are given numbers with respect to X and then calculated. Substituting the values of Table 1 and solving the equations (3a) and (3b) simultaneously, the equation (3a) becomes:

$$Y = 20.44 + 19.26b \tag{3d}$$

To test the model, the values for the first three months (from b=1 till b=3) are substituted into equation (3d) then the values for the months to be forecasted (from b= 4 till b=6) are also substituted into equation (3d).

The forecasted values are then compared with the actual values gathered from the substation as shown in Table 2. Figure 1 shows a comparison of the two values in a graphical form.

2) The Compound-Growth Model: This is another model of regression analysis. It is denoted by the formula:

$$Y = \text{antilog}(c + dX) \tag{4a}$$

Like the linear regression model, the values of c and d are also derived by solving the equations (4b) and (4c) simultaneously.

	X (month)	Y (load) MW	X^2	logY	XlogY	XY	X^3	X^4	X^2Y	Y^2
	1	41.29	1	1.6158	1.6158	41.29	1	1	41.29	1704.86
	2	55.78	4	1.7465	3.4930	111.56	8	16	223.12	3111.41
	3	79.81	9	1.9021	5.7063	239.43	27	81	718.29	6369.64
Σ (total)	6	176.88	14	5.1969	10.8151	392.28	36	98	982.7	11185.91

$$\Sigma \log Y = nc + d\Sigma X \quad (4b)$$
$$\Sigma X (\log) Y = c\Sigma X + d\Sigma X^2 \quad (4c)$$

Substituting the values of Table 1 and solving the equations (4b) and (4c) simultaneously, the equation (4a) becomes:

$$Y = \text{antilog}\,(1.4685 + 0.1432x) \quad (4d)$$

To test the model, the values for the first three months (from x=1 till x=3) are substituted into equation (4d) then the values for the months to be forecasted (from x= 4 till x=6) are also substituted into equation (4d).

The forecasted values are then compared with the actual values gathered from the substation as shown in Table 3. Figure 2 shows a comparison of the two values in a graphical form.

3) The Cubic Regression Model: This model is denoted by the formula:

$$Y = a_0 + a_1X + a_2X^2 \quad (5a)$$

The values of $a_{0,}$ a_1 *and* a_2 can be derived simultaneously using these formulas:

$$a_o n + a_1\sum x_i + a_2\sum x_i^2 = \sum y_i \quad (5b)$$
$$a_o\sum x_i + a_1\sum x_i^2 + a_2\sum x_i^3 = \sum x_i y_i \quad (5c)$$
$$a_o\sum x_i^2 + a_1\sum x_i^3 + a_2\sum x_i^4 = \sum x_i^2 y_i \quad (5d)$$

The values to be used in this calculation are all derived from Table 1. Substituting the values of Table 1 and solving the equations (5b), (5c) and (5d) simultaneously, the equation (5a) becomes:

$$Y = 36.34 + 0.18X + 4.77X^2 \quad (5e)$$

To test the model, the values for the first three months (from x=1 to x=3) are substituted into equation (5e) then the values for the months to be forecasted (from x= 4 to x=6) are also substituted into equation (5e).

The forecasted values are then compared with the actual values gathered from the substation as shown in Table 1. The MAPE and RMSE of the different models are calculated by substituting values into equations (1) and (2). The values are shown in Table 5.

Table 1 Table of values used for the computation of the respective models.

Month	Actual (MW)	Forecast (MW)
August	41.29	39.7
September	55.78	58.96
October	79.81	78.22
November	88.56	97.48
December	53.33	116.74
January	75.26	136

5. Presentation of forecast results

Table 1 above shows the computation of the respective models. The values in Y are the values of the average load consumed in each month where the first month is August 2012, second month is September 2012 and third month is October 2012.

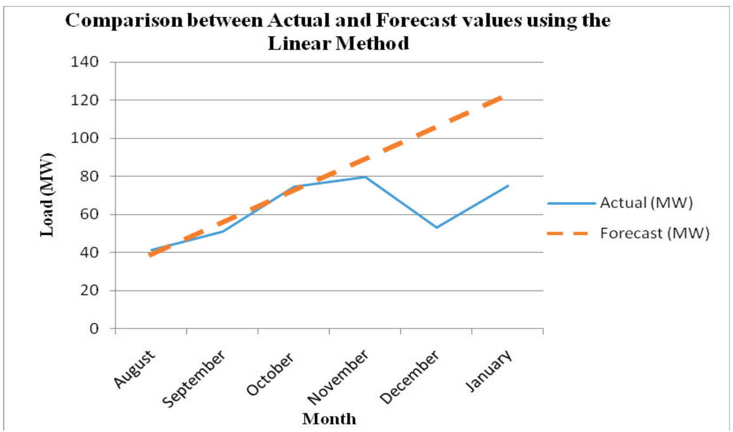

Table 2: A table showing the actual and forecasted values using the ***linear method***.

Figure 1 shows the graph that compares the actual and forecasted average load consumption of Covenant University from August 2012 – January 2013 using the linear method.

Table 3: A table showing the actual and forecasted values using the ***compound growth*** method.

Month	Actual (MW)	Forecast (MW)
August	41.29	40.9
September	55.78	56.87
October	79.81	79.09
November	88.56	109.98
December	53.33	152.93
January	75.26	212.67

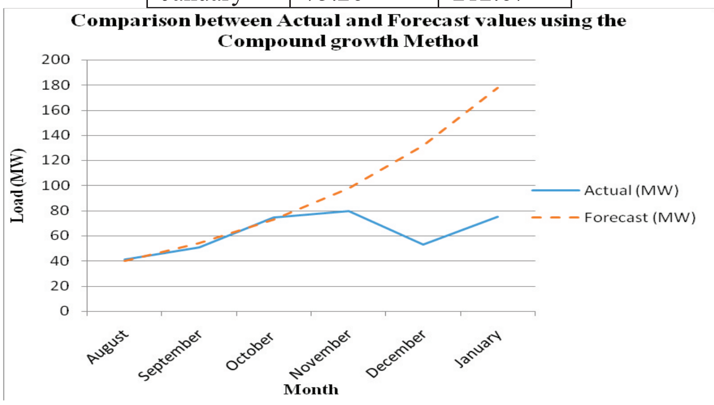

Figure 2 shows the graph that compares the actual and forecasted average load consumption of Covenant University from August 2012 – January 2013 using the compound growth method.

Table 4: A table showing the actual and forecasted values using the cubic regression method.

Month	Actual (MW)	Forecast (MW)
August	41.29	41.29
September	55.78	55.78
October	79.81	79.81
November	88.56	113.38
December	53.33	156.49
January	75.26	209.14

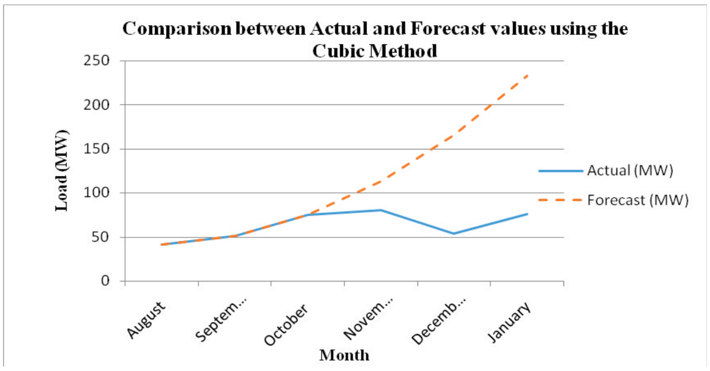

Figure 3 shows the graph that compares the actual and forecasted average load consumption of Covenant University from August 2012 – January 2013 using the cubic method growth method.

Table 5: A table showing the MPSE and RMSE values of the three regression methods.

Methods	MAPE	RMSE
Linear model	0.5792	41.34
Compound growth	1.016	75.12
Cubic model	1.536	113.16

5. Conclusion

A medium-term load forecast of an educational institution is the thrust of this paper. Regression analysis of sample load data (collected between January 2012 and January 2013) has been detailed. Results of the models used reveal a drastic decrease in load consumption in the month of December. This means that fuel supply can be at the minimum during this period, and major maintenance operation can take place as well. Meanwhile, this drop in load consumption accounts for the huge deviation in the forecast values for the three regression models. Work is under way to incorporate artificial intelligence into the study.

References

[1] http://en.wikipedia.org: 'Definition of Forecasting', Wikipedia.
[2] D.C Montgomery, L.A Johnson, and J.S Gardiner. "Forecasting & Time Series Analysis" 2nd ed. New York: McGraw Hill, 1990.
[3] Alexander Bruhns, Gilles Deurveilher and Jean-Sébastien Roy, "A Non-Linear Regression Model for Mid-term Load Forecasting and Improvements in Seasonality." 15th Power Systems Computation Conference (PSCC), Liege, 22-26 August 2005.
[4] Zaid Mohammed and Pat Bodger "Forecasting electricity consumption in New Zealand using economic and demographic variables", Christchurch, New Zealand. 2003.
[5] Abu-Shikhah N., Elkarmi F., Aloquili O., "Medium-Term Electric Load Forecasting Using Multivariable Linear and Non-Linear Regression" *Smart Grid and Renewable Energy,* 2, 2011, pp. 126-135.
[6] Eneje I.S, Fadare D.A, Simolowo O.E, Falana A., "Modelling and Forecasting Periodic Electric Load for a Metropolitan City in Nigeria, *African Research Review,* Vol. 6(1), Serial No. 24, 2012, pp. 101-115.
[7] Y. Aslan, S. Yavasca and C. Yasar, "Long Term Electric Peak Load Forecasting of Kuthaya Using Different Approaches", *International Journal on 'Technical and Physical Problems of Engineering'(IJTPE),* Iss. 7, Vol. 3, No.2, Jun. 2011, pp. 87 -91.
[8] Alexander Bruhns, Gilles Deurveilher and Jean-Sébastien Roy, "A Non-Linear Regression Model for Mid-term Load Forecasting and Improvements in Seasonality." 15th Power Systems Computation Conference (PSCC), Liege, 22-26 August 2005.
[9] Anil Karacayir, 'Short Term Electricity Load Forecasting with multiple Linear Regression and Artificial Neural Network', October 2012.
[10] Badran S.M, Abouelatta O.B, 'Forecasting Electrical Load using ANN combined with Multiple Regression Method', The Research Bulletin of Jordan ACM, Volume II (II), ISSN: 2078-7952, Page 52 – 58.
[11] Saravanan S, Kannan S and Thangaraj C, 'India's Electricity Demand Forecast Using Regression Analysis and Artificial Neural Networks based on Principal Components', ICTACT Journal On Soft Computing, Volume 02, Issue 04, July 2012, page 36 5 – 370.
[12] Oamek G.E and English B.C, "A Review of Load Forecasting Methodologies"
[13] Douglas J. Leoth, Martin Heidl and John V. Ringwood, *Gaussian Process Prior Models for Electrical Load Forecasting, 8th International Conference on Probabilistic Methods Applied to Power Systems,* Iowa State

University, Ames, Iowa, USA. Sept 12-16 2004.
[14] N. Amjady, “Short-Term Hourly Load Forecasting Using Time Series Modeling with Peak Load Estimation Capability”, IEEE Transactions on Power Systems, Vol. 16, No. 3, pp. 498 – 505, August 2001.
[15] Gang Shen, EE653 Project #1, pp. 1-10.
[16] Electric Load Forecasting: Probing the Issues with Models. EMF Report 3, Volume 1. April 1979. Energy Modelling Forum, Stanford University, Stanford, California.